INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 04 | APRIL - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Car Parking Occupancy Detection Using Deep Learning

Mani Kiran Chitirala
B.Tech Student
Department of ECE
RVR&JC College Of Engineering
Guntur,India
cheetiralamanikiran6@gmail.com

Tevendra Devarakonda
B.Tech Student
Department of ECE
RVR&JC College Of Engineering
Guntur,India
tevendradevarakonda@gmail.com

Lakshmi Venkata Harish Chaluvudi
B.Tech Student
Department of ECE
RVR&JC College Of Engineering
Guntur,India
harish.chaluvudi02@gmail.com

Guru Bhaktha Anjaneyulu Avvaru
B.Tech Student
Department of ECE
RVR&JC College Of Engineering
Guntur,India
guruavvaru53@gmail.com

Abstract—: **It can be difficult to find a parking space in a busy neighborhood because you never know when one will become available. Due to the rise of automobiles and fall in parking spots, this task even more challenging. To increase the dependability of smart parking systems, numerous researchers have sought to improve them. In smart parking systems, computer vision is a more advantageous method for detecting parking spaces than sensors. Many parking spots may be monitored by a single smart camera, which is also reasonably easy to install and maintain.**

To identify parking occupancy in camera-captured photos, a CNN model is employed. The CNN model will be trained using the CNRpark+Ext and PKLot datasets, which contain 4081 and 12,416 parking lot photos under various weather scenarios, respectively. The CNN architectures, CmAlexNet and mAlexNet, are five-layer designs built on the eight-layer AlexNet architecture, in which processing speed is increased by removing the completely connected layer and the third and fourth convolutional layers. To increase accuracy, the first convolutional layer of CmAlexNet's filter is resized from 11x11 to 13x13, and batch normalization [1] is used instead of the original method. Three key performance metrics are evaluated between CmAlexNet and mAlexNet: recall, precision, and accuracy in classification.

Keywords: Car Parking , Computer Vision , Deep learning , CNN, mAlexnet , cmAlexnet , CNRpark dataset , PKLot dataset.

I.INTRODUCTION

According to research conducted by the online pre-owned car marketplace Cars24,there are approximately 30 million cars on roads of India.In many highly populated areasthis number still increasing.With the increasing number of vehicles, there's a growing demand for smart parking management systems, particularly in areas like shopping malls and workplaces.These systems aim to optimize parking space utilization and enhance the overall parking experience for users. Most of them use sensors to monitor each individual parking slot which makes them expensive.Using computer vision for the implementation of smart parking systems with the pre-installed surveillance cameras is less expensive and cost efficient as one camera can be used monitor multiple parking.

A number of methods have recently been proposed for monitoring parking lot occupancy that rely on the use of video cameras [2][3][5]. Nevertheless, there is still an issue with detecting unoccupied parking spaces utilizing simply visual information, even with these admirable efforts. The majority of these methods rely on specific visual techniques designed for that particular situation, and therefore don't generalize well to other parking lots. In this paper, we provide a deep Convolutional Neural Network (CNN) based distributed, economical, scalable, and effective parking occupancy detection system .Deep learning, especially Convolutional Neural Networks (CNNs), has become a powerful tool in various fields such as Image recognition and classification, Natural language processing (NLP), Computer vision, Health care etc due to its ability to learn representations from complex data. These deep learning models can scale Effectively with increasing amounts of data, often leading to improved performance as more data becomes available and also are capable of generalizing well to unseen data provided they are trained on diverse and representative datasets. This ability to generalize allows models to perform effectively in real-world scenarios. Proposed model is robust to disturbances due to fractional occlusions, nearness of shadows, variety of light conditions, and shows a good generalization property.Moreover, the classification phase needs low computational resources, making it suitable for embedded environments such as Raspberry Pi.

II. RELATED WORK

Texture classification plays a crucial role in various computer vision applications such as object recognition and scene understanding. The method proposed by [4] introduces a multi-resolution approach for gray-scale and rotation invariant texture classification using local binary patterns (LBP). Their technique leverages the discrimination of sample and prototype distributions based on the occurrence histogram of local binary patterns.

Notably, the approach achieves robustness against gray-scale variations and computational simplicity due to its invariant nature and efficient implementation. Moreover, the combination of rotation invariant variance measures with LBP operators enhances the discriminatory power, demonstrating promising results in various rotation-invariant texture analysis tasks. Several studies have built upon this foundation, exploring extensions and optimizations of the LBP framework for improved texture classification performance across diverse datasets and applications. The method proposed by [2] is based on parking space detection using inter space correlation. Efficient parking space detection is essential for urban traffic management and intelligent transportation systems. The method proposed by [2] addresses this challenge by employing an 8-class Support Vector Machine (SVM) classifier with probabilistic outputs to distinguish between empty and occupied parking spaces.

Noteworthy is their incorporation of inter-space correlation using a Markov Random Field (MRF) framework, leading to improved detection performance even under challenging conditions such as occlusion and shadowing. The experimental results validate the robustness of the proposed approach in real-world scenarios, demonstrating its potential for deployment in smart parking systems and urban planning applications. Further research in this area focuses on refining the fusion strategies and optimizing computational efficiency to scale the approach for large-scale deployment in urban environments. [3] discusses about vehicle detection in satellite images using Hybrid Deep Convolutional Neural Network.

Vehicle detection in satellite images poses unique challenges due to varying scales and complex environmental conditions. [3] address this challenge by proposing a hybrid deep convolutional neural network (HDNN) that extracts variable-scale features by partitioning the maps of the last convolutional layer and max-pooling layer into multiple blocks. Their approach surpasses traditional deep neural networks (DNNs) by effectively capturing features across different scales, leading to superior performance in vehicle detection tasks. Comparative experiments demonstrate the efficacy of the HDNN in complex environments, outperforming conventional DNNs and existing feature-basedapproaches. Ongoing research in this area focuses on exploring advanced network architectures and incorporating temporal information for dynamic vehicle tracking in satellite imagery, aiming to enhance the capabilities of vehicle detection systems for surveillance and reconnaissance applications.

- 1. Multiresolution Gray-Scale and Rotation Invariant Texture Classification.
- 2. Robust Parking Space Detection Considering Inter-Space Correlation.
- 3. Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks.
- 4. Quotation marks only when a complete thought or name is cited, such as a title or full quotation.

III. DEEP LEARNING FOR CAR PARKING OCCUPANCY DETECTION

In this paper we propose a deep learning model to classify the parking lots according based on their occupancy .Deep learning is a branch of Artificial Intelligence (AI), it focuses neural networks which are inspired from the working of the human brain .These network contains multiple layers of neurons or nodes enabling it to learn complex pattern and solve complex with nearly human level accuracy.

Unlike traditional machine learning techniques that rely on handcrafted features, the deep learning algorithms hierarchical representations from the raw data enabling them to learn complex patterns.

We use a Artificial Neural Network (ANN) called as Convolutional Neural network (CNN), which are commonly used in deep learning for tasks like object detection , image recognition and natural language processing . These networks usually contain large number of layers,each of them performs a mathematical operation and produces output which is given as input to next layers.CNN processes the grid like data such as images efficiently because of its convolutional layers.Which enables it to discern the spatial correlation between the neighboring pixels which makes them better than any other neural networks.
There are two phases in building a CNN model namely training phase , which is computationally expensive and time taking and the other phase is prediction phase which is quite fast and efficient. In case of classification problem such as ours, the output is classes on which the model is trained.

Our solution is light weighted and simple which enables it to run on any embedded devices .It is better than using sensors because of its cost efficient and is easily scalable. Our solution is less expensive because a single camera can monitor multiple parking slots unlike sensors where a single sensor should be used to monitor single parking slot. Also we can increase the number of slots to be monitored by just adding a another camera. Our model is trained with two datasets namely CNRpark dataset and PKLot dataset.

IV. DATA SET

Fig .1. Samples of CNRPark data set in both status :Busy(first row) free (second row).

For training, we used the CNRPark and PKLot datasets.

PKLot [5] comprises around 700.000 photos of parking spots taken from various parking lots arranged according to the day of capture and the kind of weather (rainy, cloudy, or sunny). Fig .2 shows samples of the PKLot dataset.

CNRPark dataset developed by [6]. It was acquired by gathering over 250 photos of the parking lot over the course of many days using two different cameras that were positioned to capture diverse viewpoints, angles of view, lighting conditions, and occlusion patterns. The original screenshots are partitioned into many patches, one for each parking places.

CNR park dataset developed by [6]. It was acquired by gathering over 250 photos of the parking lot over the course of many days using two different cameras that were positioned to capture diverse viewpoints, angles of view, lighting conditions, and occlusion patterns. The original screenshots are partitioned into many patches, one for each parking places.

The size of each of these patches is a square that varies with distance from the camera; the closest patches are larger than the furthest. Next, these patches are labeled with the occupancy status of the matching parking space.

CNRPark consists of 12,584 named patches. It records various lighting circumstances, such as global or partial shadowing of cars and partial occlusions caused by obstructions like trees, lampposts, or other vehicles. This makes it possible to train a classifier that can identify the majority of scenarios that arise in real-world scenarios.Fig.1 shows the patches of both status free and occupied.

The following are CNRPark and PKLot's primary distinctions:

a) The size of PKLot is greater than that of CNRPark (695.899 vs. 12.584 total images).

b) The masks for parking spaces in CNRPark are non-rotated squares, and as a result, images frequently do not cover the entire parking space volume; in contrast, the images in PKLot are extracted using rotated rectangular masks, which are then straightened to produce a more precise coverage of the parking space.

c) In addition to being mostly obscured by trees and lampposts, CNRPark also consists of substantially occluded areas that are part of PKLot's collection of segmented spaces.

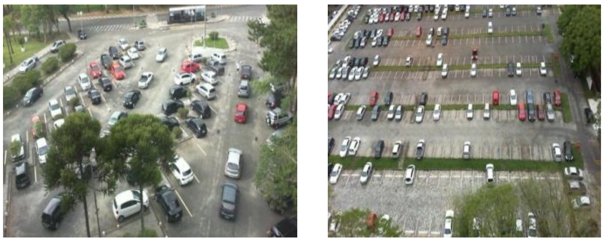

Fig .2. Samples of Pklot dataset showing two different parking lots.

Because photos are captured from PKLot's lower point of view, there are more occlusions from nearby cars. Due to these factors, PKLot's classification is less difficult than CNRPark's, which exhibits greater picture variability. By utilizing two distinct datasets, we were able to increase the ability of the model to generalize the unseen data.

V. METHODOLOGY

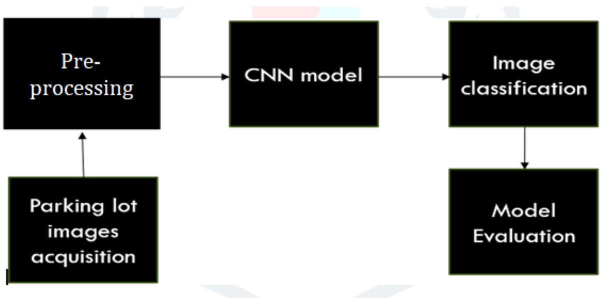

Fig .3.Methodology.

Image segmentation:

The images contain multiple parking slots, but we require the individual images of parking lots to give input to model. So, we process the images before giving input to model. The processing involves splitting the parking lot images (containing multiple parking slots) to individual parking slot images. Fig.4. shows the example of splitting the image into patches.

Fig .4. The parking lot image(right) is split into individual parking space image(left).

Re-sizing images:

The resolution of individual parking slot images is different from one another, but the input resolution of our model is constant i.e. 150 x150x3. So, we resized the images into required resolution.Fig.5. shows the example of resizing the patch into 150x150x3.

Fig .5.The image of individual parking space (right) is resized into 150x150x3(left).

mAlexnet:

We developed our model with architecture discussed in [6] shown in Fig.6, which is modified from AlexNet called mAlexnet. The Alexnet [7] consists of 5 convolutional layers and three fully connected layer Three convolutional layers (conv1 , conv2 , conv5) are followed by max pooling layers. Each layer uses Relu as the activation function except the output layer.

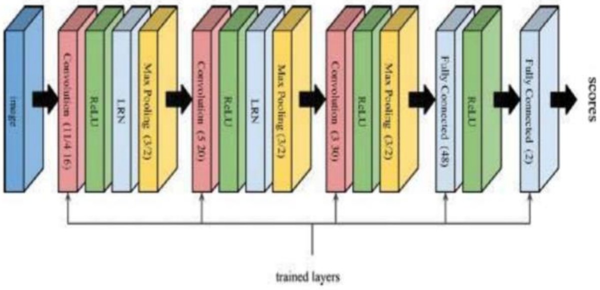

Fig.6. mAlexnet architecture

Rectified Linear Unit:

The ReLU is a activation function that introduces non linearity in neural networks, it outputs Zero if its non-positive otherwise outputs the input.

Local Response Normalization LRN:

Local response Normalization [8] is a technique used to enhance the generalization by normalizing the activation of neurons within local regions of input feature maps in CNN.

LRN encourages the completion between the adjacent neurons prevents overfitting and helps in generalizing to unseen data.

It divides each neuron's activation function by a factor that includes the sum of squares of activation functions in its neighboring neurons that are scaled by hyper parameters like alpha, beta and k.

The mAlexnet consists of only three convolutionlayers, which are followed by max pooling layers and two fully connected. All the layers are followed by ReLU activation function except the last layer which is an output layer. Theconvolution layers are followed by Local Response Normalization (LRN) except the third convolution layer(conv3). The no of layers is decreased to have less computation and obtain a light weighted model since it is a binary classification. The no of filters in conv 1-3 and neurons in fc4-5 are decreased to match dimensions of the problem. Since the number neurons are decreased in fc 4-5, no dropout regulation is adapted in fc 4-5.

cmAlexnet:

We developed our model, customized mAlexnet (cmAlexnet) by making some changes to the mAlexnet shown in Fig.7.

We changed the first size from 11X11 to 13x13 in first convolutional layer allowing it to capture more complex features and patterns and increases generalization. Also we used batch normalization instead of local response normalization because the batch normalization stabilizes the training by reducing the internal covariate shift and it also regulates improving the generalization.

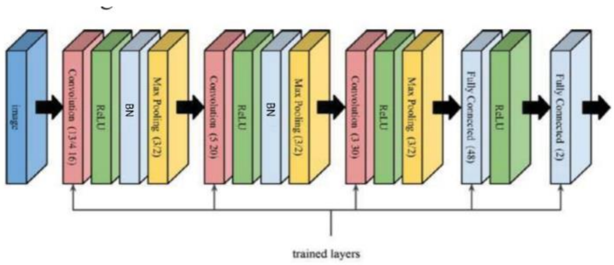

Fig.7. cmAlexnet architecture

Batch normalization:

A deep learning technique called batch normalization (BN) calculates the mean and standard deviation of activations within each mini-batch during training in order to normalized the input to each layer. Higher learning rates can be used since BN stabilizes training dynamics by reducing internal covariate shift. Moreover, BN serves as a regularizer, which lowers overfitting and enhances the model's capacity to generalize to new data. This is accomplished by employing learnable parameters to scale and shift the normalized activations, enabling the model to adaptively modify the variance and mean. All things considered, BN has established itself as a standard element of contemporary neural network topologies, facilitating faster convergence and more stable training.

BN's normalization over the entire mini-batch offers a more stable and consistent approach compared to methods like Local Response Normalization (LRN), which normalize activations based on local neighborhoods. Because of the decreased sensitivity to differences within individual samples, this improves generalization and streamlines the training process. Furthermore, BN is a recommended option in deep learning applications due to its effectiveness in a wide range of network designs and workloads. The efficiency and efficacy of deep neural networks in a variety of fields have been greatly enhanced by Batch Normalization, which also speeds up convergence and improves generalization.

TABLE I

Model	Training Accuracy	Training Loss	precision	recall	Testing Accuracy	Testing Loss
mAlexnet	94.29	16.47	87.74	96.27	94.97	15.03
Customized mAlexnet	99.16	2.39	98.73	98.76	99.24	2.65

For CNRPark dataset

Results of training the both mAlexnet and cmAlexnet onn CNR park dataset.

TABLE II

Model	Training Accuracy	Training Loss	precision	recall	Testing Accuracy	Testing Loss
mAlexnet	100	0.028	100	100	96.95	42.32
Customized mAlexnet	99.69	0.87	99.83	99.62	97.40	23.90

For PKLot dataset

Results of training the both mAlexnet and cmAlexnet onn CNR park dataset.

VI. RESULTS

This paper presents a comparative analysis of the performance of a customized variant of the mAlexnet convolutional neural network (CNN) model against its original counterpart. The evaluation is conducted on two distinct datasets: CNRPark and PKLot as mentioned earlier, both of which present challenges in vehicle detection and classification tasks.

CNRPark dataset:

In Fig.11,Fig.12. shows the training accuracy and training loss on CNRPark dataset of mAlexnet and cmAlexnet respectively. On CNRPark dataset our customized mAlext outperformed the original mAlexnet in every metric obtaining a training accuracy of 99.16 compared to 94.29 obtained by original mAlexnet. Testing accuracy and loss metrics are also significantly better than original mAlexnet and not only those precision and recall rate is also considerably better. All these values represented in Table 1.

PKLot dataset:

In Fig.9,Fig.10.shows the training accuracy and training loss on Pklot dataset of mAlexnet(first row) and cmAlexnet(second row) . On PKLot dataset however our customized mAlexnet slightly falls behind in terms of training metrics but not more than 0.5 percent nonetheless our customized model performed better in testing metrics. The lack in training performance is compensated with improved testing performance, despite this in a real-life scenario, testing metrics are generally considered more important as they reflect the model's performance on new, unseen data, which is crucial for assessing its real-world applicability. All the experimental results are available in Table 2.

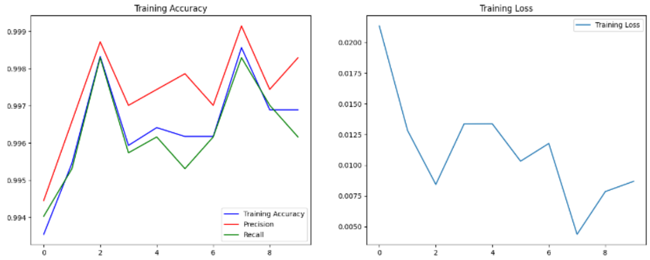

Fig.8.Training accuracy and training loss graphs of mAlexnet on PKLot dataset.

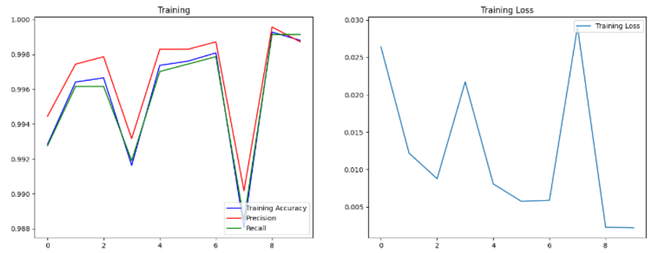

Fig.9.Training accuracy and training loss graphs of cmAlexnet on PKLot dataset.

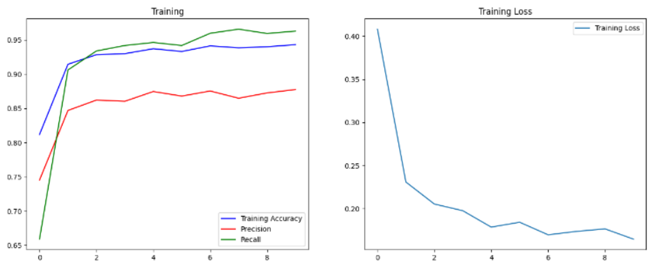

Fig.10.Training accuracy and training loss graphs of mAlexnet on CNRPark dataset.

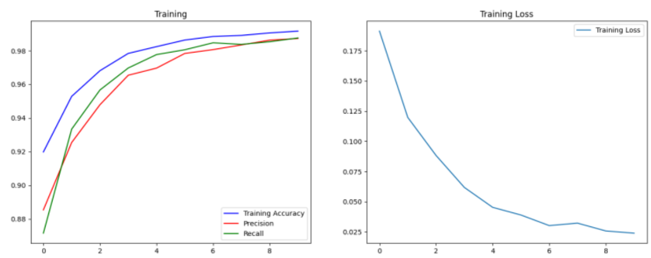

Fig.11. Training accuracy and training loss graphs of cmAlexnet on CNRPark dataset.

VII. CONCLUSION

In conclusion, the study examines how well a modified version of the mAlexnet Convolutional Neural Network (CNN) model performs in comparison to its original model when it comes to computer vision-based parking space occupancy identification. In comparison to the original model, the customized mAlexnet performed better on the CNRPark dataset across all measures, with noticeably improved training, testing, precision, and recall rates. Furthermore, the customized model surpassed the original model in testing metrics on the PKLot dataset, indicating its promise for real-world applications where testing performance is crucial, even though it fell behind in training metrics. All things considered, the tailored mAlexnet has potential for improving the dependability of smart parking systems through precise computer vision-based occupancy detection.

VIII. REFERENCES

[1] [1] V. Thakkar, S. Tewary and C. Chakraborty, "Batch Normalization in Convolutional Neural Networks — A comparative study with CIFAR-10 data," *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*, Kolkata, India, 2018, pp. 1-5, doi: 10.1109/EAIT.2018.8470438.

[2] [2] Q. Wu, C. Huang, S.-y. Wang, W.-c. Chiu, and T. Chen, "Robust parking space detection considering inter-space correlation," in Multimedia and Expo, IEEE International Conference on. IEEE, 2007, pp. 659–662.

[3] [3] C. G. del Postigo, J. Torres, and J. M. Menendez, "Vacant parking area estimation through background subtraction and transience map analysis," IET Intelligent Transport Systems, 2015.

[4] [4] N. Dan, "Parking management system and method," *Jan. 2002, US Patent App. 10/066,215.*

[5] [5] P. R. de Almeida, L. S. Oliveira, A. S. Britto, E. J. Silva, and A. L. Koerich, "Pklot–a robust dataset for parking lot classification,"*Expert Systems with Applications, vol. 42, no. 11, pp. 4937–4949, 201.*

[6] [6] G. Amato, F. Carrara, F. Falchi, C. Gennaro and C. Vairo, "Car parking occupancy detection using smart camera networks and Deep Learning," 2016 IEEE Symposium on Computers and Communication (ISCC), Messina, Italy, 2016, pp. 1212-1217, doi: 10.1109/ISCC.2016.7543901.

[7] [7] X. Chen, S. Xiang, C.-L. Liu, and C.-H. Pan, "Vehicle detection in satellite images by hybrid deep convolutional neural networks," Geoscience and Remote Sensing Letters, IEEE, vol. 11, no. 10, pp. 1797–1801, 2014.

[8] [8] Kangbae Lee, Sang Ha Sung, Doo-hwan Kim, Sung-ho Park "Verification of Normalization Effects Through Comparison of CNN Models", 2019 International Conference on Multimedia Analysis and Pattern Recognition (MAPR),doi: 10.1109/MAPR.2019.8743531.

[9] [9] T. Ojala, M. Pietikainen, and T. M ̈ aenp ̈ a ̈a, "Multiresolution gray-scale ̈ and rotation invariant texture classification with local binary patterns," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, no. 7, pp. 971–987, 2002.

[10] [10] Q. Wu, C. Huang, S.-y. Wang, W.-c. Chiu, and T. Chen, "Robust parking space detection considering inter-space correlation," in Multimedia and Expo, IEEE International Conference on. IEEE, 2007, pp. 659–662.

[11] [11] R. Yusnita, N. Fariza, and B. Norazwinawati, "Intelligent Parking Space Detection System Based on Image Processing," Journal of Innovation, Management and Technology, vol. 3, no. 3, pp. 232–235, 2012.

[12] [12] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal Loss for Dense Object Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 318-327, 2020

[13] [13] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real- Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 201T.

[14] [14]Moranduzzo and F. Melgani, "Automatic car counting method for unmanned aerial vehicle images," IEEE Transactions on Geoscience and remote sensing, 52(3), 1635-1647, 2013.

[15] [15] T. Moranduzzo and F. Melgani, "Detecting cars in UAV images with a catalog-based approach," IEEE Transactions on Geoscience and remote sensing, 52(10), 6356-6367, 2014

[16] [16] D.G. Lowe, distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60(2), pp. 91-110, 2004.